WYC:dks 5/13/96 3382-40231app

PATENT

MAIL ROOM
MAY 17 1996
PAT. & TR. OFF.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of

Vanderwende et al.

Application No. 08/227,247

Filed: April 13, 1994

For: METHOD AND SYSTEM FOR COMPILING A LEXICAL KNOWLEDGE BASE USING BACKWARDS-LINKING NATURAL LANGUAGE PROCESSING (As Amended)

Examiner: R. Weinhardt

Date: May 13, 1996

Art Unit 2411

CERTIFICATE OF MAILING

I HEREBY CERTIFY THAT THIS PAPER AND THE DOCUMENTS REFERRED TO AS BEING ATTACHED OR ENCLOSED HEREWITH ARE BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE ON May 13, 1996 AS FIRST CLASS MAIL IN AN ENVELOPE ADDRESSED TO: ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, DC 20231.

William Y. Conwell
Attorney for Applicant

TRANSMITTAL LETTER

96 MAY 21 AM 10:28

BOX AF
ASSISTANT COMMISSIONER FOR PATENTS
Washington, DC 20231

Sir:

Enclosed for filing in the above-referenced matter are the following:

Appeal Brief (37 pages) (in triplicate)

Our check for $290 to cover filing of Appeal Brief.

Please charge any additional fees which may be required in connection with filing this amendment and any extension of time, or credit any overpayment, to Deposit Account No. 02-4550. A **duplicate** copy of this sheet is enclosed.

Respectfully submitted,

KLARQUIST SPARKMAN CAMPBELL LEIGH & WHINSTON, LLP

By__________________
William Y. Conwell
Registration No. 31,943

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391

cc: Katherine Welty

BOX AF 2900 120 1073 G240

WYC:dks 5/13/96 3382-40231app.brf

PATENT

MAILROOM MAY 16 1996 PAT. & TRADEMARK OFF.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

G. Hesper #13 Appeal Brief 5/29/96

In re application of

Vanderwende et al.

Application No. 08/227,247

Filed: April 13, 1994

For: METHOD AND SYSTEM FOR COMPILING A LEXICAL KNOWLEDGE BASE USING BACKWARDS-LINKING NATURAL LANGUAGE PROCESSING (As Amended)

Examiner: R. Weinhardt

Date: May 13, 1996

Art Unit 2411

CERTIFICATE OF MAILING

I HEREBY CERTIFY THAT THIS PAPER AND THE DOCUMENTS REFERRED TO AS BEING ATTACHED OR ENCLOSED HEREWITH ARE BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE ON May 13, 1996 AS FIRST CLASS MAIL IN AN ENVELOPE ADDRESSED TO: ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, DC 20231.

William Y. Conwell

William Y. Conwell
Attorney for Applicant

96 MAY 21 AM 10: 28 GROUP 2400

APPEAL BRIEF

BOX AF
ASSISTANT COMMISSIONER FOR PATENTS
Washington, DC 20231

Sir:

This brief is in furtherance of the Notice of Appeal filed March 11, 1996.

The fee required under 37 CFR 1.17(f) is enclosed.

300 SL 05/21/96 08227247
1 120 290.00 CK

CONTENTS

I. REAL PARTY IN INTEREST

The real party in interest is Microsoft Corporation, by an assignment from the inventors recorded at Reel 6964, Frame 337.

II. RELATED APPEALS AND INTERFERENCES

None.

III. STATUS OF CLAIMS

Claims 20-21, 24-30, 33-43 and 45 are finally rejected and appealed. Claims 16-19 are allowed, and claims 22-23, 31-32 and 44 are objected-to. Claims 1-15 are cancelled.

IV. STATUS OF AMENDMENTS

An Amendment After Final was filed on February 12, 1996, and overcame rejections of the claims under § 112, ¶2.

The Amendment After Final also sought to amend the application Title to overcome an objection thereto. Although not particularly stated in the Advisory Action (paper 9), applicants presume the amendment to the title was entered and overcame the objection.

V. SUMMARY OF THE INVENTION

A. Background

We take common sense for granted. Computers don't enjoy this advantage. Everything a computer "knows" it must be told, or it must deduce from what it already knows.

For nearly fifty years, futurists have predicted that machines emulating human reasoning are just around the corner. Yet that corner still hasn't been turned. The reason? The unavailability of a repository of common-sense information from which computers can draw understanding of our world.

Consider the pair of sentences:[1]

The ink is in the pen.

The pig is in the pen.

The first sentence's "pen" is a writing implement, the second's is a pig-sty or corral. How do you know that? It's not English that tells you, it's your common sense knowledge.

Consider the following example:

- *The little girl saw the bicycle in the window. She wanted it.*
- *The little girl saw the bicycle in the window. She pressed her nose up against it.*

What does the word "it" refer to in each of these sentence pairs, the bicycle or the window? How do you know that? You know because of your knowledge of

[1] Examples taken from Lenat, *Building Large Knowledge-Based Systems* (1989).

the real world, of human emotions and capabilities and anatomy, of physical and mental and physiological limitations, and so on -- *not* because of English grammar and syntax.

The need for a repository of general information that would enable a computer to make these distinctions (a depository termed a "lexical knowledge base" by artisans in the artificial intelligence (AI) field) has long been recognized. Efforts to produce such knowledge bases have existed for decades, including the massive Cyc Project discussed below.

An approach that was originally thought to hold great promise is "Natural Language Understanding" (NLU). According to this theory, a computer could be provided with a large textual source of information, such as a dictionary, and could analyze the contents (with preprogrammed knowledge of English syntax and linguistics) to extract millions of snippets of information. (Artisans term such snippets "assertions" or "rules.") The extracted information, collectively, would represent a large body of such common sense knowledge.

In practice, Natural Language Understanding approaches to compiling knowledge bases have yielded disappointing results. A reason is that such definitions, themselves, are surprisingly sparse sources of information. Consider a dictionary definition for "flower," which is noteworthy more for the information it omits than for what it provides:

> **flower**: the part of a plant, often beautiful and colored, that produces seeds or fruit.

Missing from this definition is any detailed description of the physical structure of flowers, information about what kinds of plants have flowers, and so on. Even the important fact that flowers prototypically have a pleasant scent goes unmentioned.

A computer's "understanding" of "flower" can be augmented by exploring definitions of words used in the definition of flower ("plant," "seeds," etc.), but such strategies (so-called "forward linking" strategies) are relatively unproductive.

Consequently, the prior art has widely rejected Natural Language Understanding as an unattainable "Free Lunch."

B. The Invention

The present invention relies on a different approach to augmenting the sparse information available from dictionaries and the like -- one that finally makes Natural Language Understanding a viable way to *automatically* compile massive knowledge bases. That technique is "backward linking."

In compiling information associated with the word "flower," systems according to most of the pending claims do not rely just on the definition of "flower," nor on the definitions of terms used in the definition of "flower." Instead, such traditional definitional information is here augmented by information gleaned from *other* dictionary entries which use the term "flower" in *their* definitions. "Petal," "garden" and "rose" are but some of the rich variety of terms that are identified in this fashion. The contexts in which these associated terms are found

can then be analyzed to discern semantic relationships between each of these terms and the original word "flower."

More particularly, the knowledge base entry for "flower" initially comprises just the information gleaned from the definition of "flower."[2] Likewise for all other words in the dictionary. But as each definition in the dictionary is examined and used to compile "knowledge" on the defined term, it is also "inverted." The inverted definition is analyzed to determine how it might augment the knowledge base's knowledge of *other* terms.

Consider what happens when the system encounters the definition for "corolla:"

corolla: part of a flower

The system will use this information for its understanding of "corolla." But it *also inverts* this definition to yield:

flower: part corolla[3]

This latter snippet of information is used to augment the knowledge base entry for flower.

(The foregoing simple definition is a rare case; real-world definitions are much more complex. The "inversion" of such definitions are likewise much more complex, as illustrated by Figs. 2 and 3 of the specification, which considers the

[2] Technically, this may not actually be the *first* step.

[3] "part of" and "part" are two common "semantic relations" -- a linguistic term expressing the relationship between two words. Other common semantic relations include "purpose," "purpose of," "location," "location of" and "hypernym."

definition "**market**: a building, square, or open place where people meet to buy and sell goods.")

By applicants' automated use of inversion, the knowledge base entry for "flower" can include not just listings of flower parts, or listings of flower types, but more diverse knowledge, such as: that bees collect nectar from them, that they can be put into a vase, that they are sold from a shop by a florist, that they are rolled up until they open, etc., etc.

As noted, the prior art has widely rejected Natural Language Understanding as an unattainable "Free Lunch." The Lenat book cited in each of the Examiner's rejections is one of many prior art works expressing this view. (Lenat's approach, and the one adopted by the cited Cyc Project, is to rely on *humans* -- rather than computers -- to extract knowledge from textual sources and compile a knowledge base therefrom.) Defying this conventional wisdom, applicants' incorporation of backward linking finally makes the Natural Language Understanding a workable approach, and permits massive and comprehensive knowledge bases to be assembled quickly. (The Cyc Project, in contrast, is now beginning its second decade of work in manually compiling its knowledge base.)

Turning now more particularly to the combinations defined by the rejected claims, in one aspect (claim 20), the invention is a method for generating a lexical knowledge base in a machine. The method includes:

(a) using a natural language parser associated with the machine to parse a segment of text to obtain a logical form;[4]

(b) using the machine to extract from the logical form a semantic relation structure including at least a headword [e.g. "flower"],[5] a semantic relation [e.g. "*part_of*"], and a value [e.g. "plant"];[6]

(c) storing, in a memory associated with the machine, the semantic relation structure in association with the headword in the lexical knowledge base [e.g. "flower: *part_of* plant"];[7] and

(d) augmenting the lexical knowledge base by (1) inverting the semantic relation structure, and (2) storing, in association with the value [e.g. "plant"], the inverted semantic relation structure [e.g. "...*part* flower"].[8]

According to another aspect (claim 30), the invention is a machine-implemented method for generating a lexical knowledge base. The method includes:

(a) using a computer to automatically process a text and identify at least a first *four-element* semantic relation structure therein;[9]

---

[4] *See, e.g.*, page 8, lines 5-7; Figs. 1, 2.

[5] Bracketed text indicates examples and does not limit the invention's scope.

[6] *See, e.g.*, page 8, lines 18-25; page 11, lines 11-14; Figs. 1, 3.

[7] *See, e.g.*, page 9, lines 1-2; Fig. 1.

[8] *See, e.g.*, page 9, lines 12-23; Figs. 1, 4.

[9] *See, e.g.*, page 8, lines 5-7, 18-28; Figs. 1, 3.

(b) using the computer to invert the first semantic relation structure to yield a second, corresponding, inverted semantic relation structure;[10] and

(c) storing data from the second semantic relation structure as part of the lexical knowledge base.[11]

According to another aspect (claim 43), the invention is a system comprising:

(a) machine-implemented means for parsing a segment of text to obtain a logical form;[12]

(b) machine-implemented means for extracting from the logical form a semantic relation structure, the semantic relation structure including at least a headword, a semantic relation, and a value;[13]

(c) computer memory defining a lexical knowledge base in which the semantic relation structure associated with the headword is stored;[14] and

(d) means for augmenting the lexical knowledge base including means for:

(1) inverting the semantic relation structure;[15] and

---

[10] *See, e.g.*, page 9, lines 12-14; Figs. 1, 4.

[11] *See, e.g.*, page 9, lines 16-23; Fig. 1.

[12] *See, e.g.*, page 8, lines 5-7; Figs. 1, 2.

[13] *See, e.g.*, page 8, line 18, 22-24; Figs. 1, 3.

[14] *See, e.g.*, Fig. 8, element 104; page 9, lines 1-2.

[15] *See, e.g.*, page 9, lines 12-14, Figs. 1, 4.

(2) storing, in association with the value, the inverted semantic relation structure.[16]

While all of the foregoing aspects of the invention make use of machine inversion of a semantic relation structure, another aspect of the invention does not. According to this aspect of the invention (claim 41), the invention is an improvement to a prior art machine-implemented method of generating a lexical knowledge base by using a computer to parse a collection of texts to identify semantic relation structures, and storing data from the semantic relation structures in a lexical knowledge base. The claimed improvement comprises:

(a) using the computer to iteratively parse the same collection of texts;[17] and

(b) for each successive parsing, relying on successively enhanced versions of the lexical knowledge base;

wherein semantic relations identified in one parsing operation serve to enhance discernment of semantic relations in subsequent parsing operations.[18]

Various features of certain of the claims are not addressed in the foregoing summary; these are detailed in the discussions that follow.

---

[16] *See, e.g.*, page 9, lines 16-17, Fig. 1.

[17] *See, e.g.*, page 8, lines 5-7; page 9, lines 25-26; page 19, line 22 through page 20, line 15; Fig. 1.

[18] *See, e.g.*, page 9, lines 25-28; page 19, line 22 through page 20, line 15.

VI. ISSUES

1. Whether the Office established a *prima facie* case of obviousness as to each of the rejected claims.

2. If a *prima facie* case of obviousness was established, whether the Office improperly ignored repeated and adamant prior art teachings directly *away* from the claimed combinations.

VII. GROUPING OF CLAIMS

Each of the claims is independently patentable. The reasons therefor are detailed below, and summarized in Section IX.

VIII. ARGUMENT

The rejections must be reversed because the Office failed to establish a *prima facie* case of obviousness for any of the appealed claims.

Moreover, the Office acted improperly in ignoring explicit teachings in the cited Lenat reference which would have led an artisan to *avoid,* rather than adopt, applicants' claimed combinations.

A. The Art

1. So-Called "Applicants' Admitted Art"

As support for the final Section 103 rejection, the Office cited the Section 103 rejection expressed a year earlier in Paper 4 (mailed 1/19/95). That earlier

action cited applicants' specification, particularly "page 2, lines 9+, and page 10, lines 8+,"[19] as the *principal* § 103 reference. The action summarized the cited disclosure as follows:

> Applicant's admitted prior art teach the parsing and extraction of semantic relation structures including a headword, a relation and a value from dictionaries and their storage in association with the headword for use as a knowledge base for natural language processing.

Not so.

First, the cited pages of applicants' specification reviews the teachings of more than a dozen diverse research efforts.[20] The Examiner's unrevealed piecing together of different elements from these diverse efforts and his distilling of same into a single quoted "teaching" betrays striking hindsight.[21]

Moreover, page 9 of applicants' specification (between the two pages cited in the Action) makes clear that the Examiner's distillation is *incorrect*. There applicants noted:

> So far, it should be noted that no backward linking has occurred. Instead, the steps have only been concerned with the original dictionary definition for the word "flower." Further, it should be noted that these steps, per se, are well known in the art, being found in many

---

[19] Paper 4, ¶ 7.

[20] The Office's citation to "page 2, lines 9+," etc., renders uncertain exactly where the asserted prior art is found, so the number of research efforts reviewed is likewise uncertain.

[21] Applicants' selection of this particular grouping of research efforts is not, of course, itself prior art, and cannot be used to suggest that it would have been obvious to one of ordinary skill in the art to have so grouped the existing research and prior art.

> other NLP analysis systems (*albeit not in the context of extracting semantic relations from the logical form of definitions in on-line dictionaries*).

The italicized text directly refutes the Examiner's distillation of "admitted prior art."

Since the principal reference consists of the Examiner's selective piecing together of a large body of diverse art without any suggestion to do so -- a piecing together which directly contradicts applicants' caveat about the scope of the art -- it cannot serve as "admitted prior art" to sustain a rejection under § 103. Accordingly, the rejections premised thereon (i.e. *all* the pending rejections) must be reversed.

2. Lenat

a. The Cyc Project

About a decade ago, Lenat and Guha set out to compile an enormous database of "common sense" knowledge. As indicated above, such a database has long been viewed as the holy grail of artificial intelligence; with it, computers would be able to make sense of humans and our communications.[22]

Lenat and Guha are regarded as preeminent in this field, and their proposal quickly attracted funding. For the past ten years, they have led a massive project to compile such a database. Through the course of their work, they published widely; the Lenat book cited in the final rejection is one of many works detailing their project and its progress.

Applicants' invention is directed to the same task: the compilation of massive "common sense" databases. Applicants' approach, however, is directly opposite Lenat's.

---

[22] An early writing evidencing this long felt need is Mortimer Taub's 1961 cybernetic classic *Computers & Common Sense: The Myth of Thinking Machines*, which viewed such common sense knowledge as the missing link precluding computers from replacing human reasoning. (While skeptical that such computer knowledge could ever be achieved, Taub still counselled "constant effort to discover, to abstract, and to automate all formal elements of decision processes..." (pp. 97-98).)

Still earlier is Edmund Berkeley's 1950 treatise, *Giant Brains, or Machines That Think*, which predicted a great variety of cybernetic wonders, all premised on the hope that computers could someday be endowed with something approaching human-like "knowledge."

b. Lenat Explicitly Rejected Applicants' General Approach; It Was Error for the Office to Ignore Same

Lenat considered and analyzed various techniques that might be used to compile such a "common sense" database; his analysis is recounted in various of his writings. The approach he finally followed was to rely on human coders to interpret human knowledge (as expressed in various natural language sources, such as articles, comics, etc.), and convert it into a form suitable for entry in a computer database.

Applicants' approach, in contrast, does not rely on human coders. Instead, applicants' approach relies on a computer to interpret natural language texts (most particularly dictionaries), extract knowledge therefrom, and automatically compile a resulting database.

The general technique employed by applicants, i.e. the automated approach, *was explicitly considered, and rejected*, by Lenat.

His rejection of this approach was first explained in the cited Lenat book (1989) under the heading "No Free Lunch." There he recounted unsatisfactory experiences with prior automated techniques dating back over two decades and concluded automated compilation of knowledge bases should be avoided:

> **Hard Work (No Free Lunch)** The limited success we had with automatic program synthesis from examples and dialogues in the early seventies led us to the AM research (automated discovery of domain concepts). Its limited success led us to Eurisko, which tried to discover new heuristics as it went along. *Its* limited success led us to believe that there is no free lunch; that is, that we had to apply the tactic of last resort - hard

> work - and thus Cyc was born. We are building the needed [knowledge base] manually, one piece at a time, at least up to the crossover point where natural language understanding begins to be a more effective way of further enlarging it.[23]

In the Final Rejection, the Office ignored this contrary teaching, arguing that Lenat's teaching of "no free lunch" is an outdated 1970s-era assessment of the world -- one superseded by advancement of technology in the past two decades:

> Applicant attempts to characterize Lenat as teaching away from the combination claimed. However, the citation of Lenat used as support for this argument discussed failures in "the early seventies" and indicated that automation would not be shunned when "natural language understanding begins to be a more effective way of further enlarging" the knowledge base. Thus, Lenat is not discarding the automated approach as asserted by applicant. *As natural language understanding has certainly advanced since the date of publication of Lenat, Lenat's assumptions concerning the abilities in the art cannot be taken as the last word.*[24]

Again, not so.

While Lenat's "no free lunch" teaching may have first arisen out of his 1970s work, it is a view he still held when his cited book was published (1989), *and is a view he still holds today.*

---

[23] Lenat, *Building Large Knowledge-Based Systems*, 1989, page 26.

[24] Paper 7, page 5, ¶ 2 (emphasis added).

In 1990, Lenat/Guha published an interim report on Cyc.[25] On page 3 of this report, Lenat expressed again his "no free lunch" view, as follows:

> Here are three notable candidates we've had to reluctantly rule out (as complete "free lunches"):
> • Natural Language Understanding (NLU) is one tantalizing "free lunch" route: get a program to read English, say, then use it to rapidly build up the large [Knowledge Base] (semi)-automatically.

In 1991, Lenat/Guha published a journal article.[26] On the opening page they begin with the stark "truth:"

> Perhaps the hardest truth to face -- one that AI has been trying to wriggle out of for several decades -- is that there's probably no elegant, effortless way to *obtain* that immense knowledge base. Rather, the bulk of the effort must -- at least initially -- be manual entry of assertion after assertion.

(Emphasis in original.)

It should be recognized that applicants' work, in contrast, contravenes this "truth" and fulfills this long sought-after goal in artificial intelligence. There *is* an elegant, effortless way to obtain an immense knowledge base -- it is the invention presently claimed. Such pioneering work, *directly contrary* to established teaching, was not an obvious approach to solution of this longstanding problem.[27]

---

[25] Guha and Lenat, *Cyc: A Midterm Report*, MCC Technical Report ACT-CYC-134-90, 1990 (attached as Exhibit A).

[26] Guha and Lenat, *Cyc: A Mid-Term Report*, Applied Artificial Intelligence, 5:45-86, 1991 (attached as Exhibit B).

[27] One important indicia of non-obviousness recognized by the Supreme Court in *U.S. v. Adams*, 383 U.S. 39 (1966), is teaching away from the claimed invention by the prior art. Moreover, the Federal Circuit on several occasions has followed this precedent, holding that

Just last year, Lenat continued to argue that the "obvious" approach of machine natural language understanding should be avoided:[28]

> Why *manually* construct it, rule by rule, rather than drawing on one or both of the obvious methods for building it up automatically and painlessly: machine learning (discovering new information automatically) and/or natural language understanding (gleaning information from already-written texts)? *Very reluctantly, we concluded that in each case it would be premature to rely on those techniques, because they themselves required this selfsame large knowledge base in order to succeed.*[29]

The foregoing makes evident that Lenat's "No Free Lunch" teaching is not a relic of the 1970s, but is Lenat's current teaching to artisans in the field. It was error for the Office to disregard same.

---

teaching away from the art is a *per se* demonstration of lack of *prima facie* obviousness. *In re Dow Chemical*, 837 F.2d 469 (Fed. Cir. 1988).

[28] Lenat, D.B., "Steps to Sharing Knowledge," Proceedings Second International Conference on Building and Sharing Very Large-Scale Knowledge Bases, April 10-13, 1995, pp. 3-6 (attached as Exhibit C).

[29] Page 5, top of col. 1, emphasis added.

c. The Office Also Erred in Finding Lenat Taught That Passage of Time, Alone, Would Make Automated Approaches Viable

The Office cited Lenat's book as teaching manual building of knowledge bases *only until "natural language understanding begins to be a more effective way of further enlarging [the knowledge base]."* The Office concluded therefrom that applicants' approach isn't shunned by Lenat; only avoided until the art advances.[30]

Again, not so.

Lenat's 1995 writing makes clear what Lenat meant by the cited phrase, namely: natural language understanding won't be effective *until it can draw on a huge pre-existing knowledge base.*

In the 1995 writing quoted above, Lenat taught that it was "premature" (in *1995*) to rely on natural language understanding because it requires "this selfsame knowledge base in order to succeed."

By 1995, the Cyc project:

- had been in process for ten years;
- had generated "the first million or two rules;" and
- had consumed *a person-century of effort.*[31]

Yet as of the article's writing, Lenat *still* had not progressed to the point of using natural language understanding; that he planned for the *following* decade.

---

[30] Paper 7, page 5, ¶2, quoted *supra.*

[31] Lenat, "Steps to Sharing Knowledge," *supra.*

Applicants' invention, in contrast, requires no such prerequisite. It can automatically compile fundamental knowledge bases of the sort which Lenat counsels must be manually constructed.[32]

From the foregoing, it is evident that, even *after* the present application was filed, the art was *still* not sufficiently advanced for Lenat and followers of his teachings to pursue the approach perfected and particularly here claimed by applicants.

d. Inversion

Lenat knew of inversion; it is a fundamental linguistic operation dating back decades, and Lenat saw that it could be put to use in compiling the Cyc database.

Lenat, however, failed to recognize the tremendous power this technique could bring when applied in conjunction with the other elements of the presently claimed combinations. Instead, he treated it as just another technique employed to augment a hand-compiled knowledge base, rather than the key which would enable the long-sought[33] goal of *automatic* compilation of huge knowledge bases by natural language understanding.

---

[32] Once a knowledge base has been built, the knowledge thereby gained can be used to assist further processing of natural language sources, as discussed in applicants' specification. But the existence of a huge knowledge base is not a condition precedent in applicants' system, as it is in Lenat's teachings.

[33] In the 1990 report, Lenat noted that the quest for an elegant, effortless way to obtain an immense knowledge base had been continuing for 34 years. Page 3.

B. The Office Failed to Establish a Prima Facie Case Under § 103

The foregoing discussion has already noted several fatal flaws in the final rejection. Each of these leads the rejection to fall short of meeting the Office's *prima facie* burden:

- The principal § 103 reference is the Examiner's selective piecing together of a large body of diverse art without any suggestion to do so (and without any explanation by the Examiner of his logic therefor) -- a piecing together which directly contradicts applicants' explicit caveat about the scope of the art. Such an errant distillation is not "admitted prior art."
- The cited art adamantly counsels *against* the class of techniques followed by applicants.

The rejection is flawed in other respects too, including a failure to show the requisite "reasonable expectation of success;" consideration of only the "gist" of the invention (while impermissibly ignoring numerous claim limitations); and failure of the art to suggest the proposed combinations.

1. No "Reasonable Expectation of Success"

One of the requirements of a *prima facie* case is "a reasonable expectation of success."[34] Given Lenat's clear and repeated teachings that automatic natural

[34] MPEP § 2142.

language processing is an unworkable "free lunch," this legal requirement is missing and the rejection fails.

2. "Gist" and Limitations Ignored in Claims 21, 24, 25, 27, 28, 29, 30, 33, 34, 35, 38, 39, 40, 42 and 45

The MPEP cautions examiners against considering the "gist" of an invention.[35] Yet this is precisely what has been done in Paper 7; numerous limitations of applicants' claims are nowhere discussed or even acknowledged.

Applicants do not argue that the neglected claim limitations are all new *per se*. Rather, applicants' quandary is that the Office has failed to cite particular art showing particular limitations, and has thereby denied applicants the opportunity to explain why a particular combination of cited references thereby proposed is non-obvious. Instead, the Office presents applicants a phantom aggregation of elements, assembled from unknown contexts, combined according to the Examiner's personal view of "common sense." Such a phantasmic foil is not the reasoned combination of art required to establish a *prima facie* case.

To illustrate, the Office failed to acknowledge the following elements of applicants' claims:

- the "at least one element in addition to a headword, a semantic relation, and a value" in a semantic relation structure to be inverted (claim 21);

---

[35] MPEP § 2141.02.

- the "broad coverage parser" (claims 24, 45);
- "applying a first set of rules to a segment of text to yield a syntactic structure corresponding thereto," and "applying a second set of rules to the syntactic structure to produce the logical form" (claim 25);
- discerning different sets of semantic relations in first and second passes of a two-pass extraction process (claims 27, 38);
- extracting the semantic relation "hypernym" only in the first of a two-pass extraction process (claims 28, 39);
- extracting the semantic relations "user," "domain" and "manner" only in the first of a two-pass extraction process (claims 29, 40);
- the automatic computer processing of a text to identify a first semantic relation structure that includes at least four elements, and its subsequent automatic inversion to yield a second semantic relation structure (claim 30);
- the automatic parsing of a text by a natural language parser by "applying a first set of rules to the text to yield a syntactic structure corresponding thereto," "applying a second set of rules to the syntactic structure to produce a corresponding logical form" and "extracting from the logical form said first semantic relation structure, said first semantic relation structure including at least a headword, a semantic relation, and a value" (claim 33);

- the analysis, by a natural language parser, for at least 8 semantic relations chosen from part, part_of, typical_subject, typical_subject_of, typical_object, typical_object_of, purpose, purpose_of, location_of, located_at, and synonym (claim 34);

- string searching the text to discern semantic relation structures (claim 35); and

- "applying a first set of rules to a collection of texts to yield a collection of syntactic structures corresponding thereto," "applying a second set of rules to the syntactic structures to produce corresponding logical forms" and "applying a third set of rules to the logical forms to obtain semantic relation structures..." (claim 42);

The failure of the Office to acknowledge any of the foregoing elements, or address same during prosecution, requires reversal of the obviousness rejections of claims 21, 24, 25, 27, 28, 29, 30, 33, 34, 35, 38, 39, 40, 42 and 45, independently of any other claim(s).

3. Certain of the Office's Assertions Concerning Prior Art Are Wrong

The foregoing section detailed claims whose limitations were nowhere acknowledged during prosecution.

For other claims, the Office acknowledged certain limitations, and attempted to address same. Often, the Office made errors of fact in such analyses.

First, as noted above, all of the rejections are premised on so-called "admitted prior art." But as earlier detailed, this is an errant distillation by the Examiner, and not "admitted prior art." For this reason, alone, all of the rejections must be reversed.

Elsewhere, the Lenat reference is misread. For example, the rejection states:

> Lenat also specifically teaches that [inversion] provides an accessing advantage when using the knowledge base and provides the possible introduction of additional inferences. See section 3.3.1 on pages 83-84 of Lenat.[36]

Again, not so. The cited excerpt nowhere "specifically teaches" that inversion provides possible introduction of additional inferences. Rather, the excerpt merely notes that adding inverses to the knowledge base speeds searching and facilitates removal of unwanted entries.[37]

The rejections of applicants' claims directed to iterative processing (e.g. claims 26, 37, 41) are premised on this non-existent "specific teaching"[38] and must therefore be reversed, independently of any other claim(s).

(Moreover, such claimed methods perform iterative processing so that knowledge accumulated in the first iteration may be used to enhance the parsing, or the discernment of semantic relations, performed in the second iteration. Lenat, in

---

[36] Paper 4, page 8.

[37] Lenat at p. 84: "in order *not* to have to search the whole KB to answer this question, we need an inverse pointer..." and "If we decide to kill (obliterate) the Bill unit, we would like a way to quickly find all the places in the KB that refer to it..."

[38] Paper 4, page 8 ("Concerning the recited repetition...").

contrast, does not perform machine parsing or discernment of semantic relations; he relies on human coders. Since there is no machine parsing, or discernment of semantic relations, in Lenat that may be enhanced by a larger knowledge base, there is no reason to iteratively process.)

4. The Office Failed to Cite a Sufficient Suggestion in the Art Leading to the Proposed Combinations

The Final rejection relies on "common knowledge and common sense," *without any specific hint or suggestion in a cited reference*, as a rationale leading an artisan to modify and combine the applicants' "admitted art" and Lenat to yield each of applicants' claims 20-21, 24-30, 33-43 and 45.

The view that "the conclusion of obviousness may be made from common knowledge and common sense of a person of ordinary skill in the art *without any specific hint or suggestion in a particular reference,*" (a view first expressed in *In re Boznek*, 163 USPQ 545 (CCPA 1969)) is waning in popularity.

To avoid impermissible use of hindsight, the Federal Circuit increasingly looks for a citation to a prior art reference for a suggestion leading to the proposed combination/modification. Indeed, the Federal Circuit has never cited *In re Boznek* in any of its published opinions. More representative of the Court's current view is *In re Rijckaert*, 28 USPQ2d 1955, 1957 (Fed. Cir. 1993), which held:

> [W]hen the PTO asserts that there is an explicit or implicit teaching or suggestion in the prior art, it must indicate where such a teaching or suggestion appears...

Accordingly, appellants respectfully submit that the Examiner's basis for these rejections is legally inadequate, and rejections premised thereon should be reversed.

Moreover, applicants respectfully submit that if "common sense" can be used as the sole basis for obviousness, its application to Lenat's teachings would lead an artisan *away from*, not towards, applicants' claimed combinations.

Each of applicants' rejected claims requires the *automatic* processing of a text to extract a semantic relation structure therefrom. Lenat repeatedly and expressly instructed artisans to *avoid* such automated approaches. As such, neither common sense nor the cited art can provide the requisite suggestions leading to these claimed combinations. Absent such suggestions, the rejections are legally insufficient and must be reversed.

IX. NOTES ON INDEPENDENT PATENTABILITY OF CLAIMS

Reasons for the allowability of each of claims 21, 24, 25, 27, 28, 29, 30, 33, 34, 35, 38, 39, 40, 42 and 45 -- independently of any other claim(s) -- were detailed in Section VIII.B.2, above.

Reasons for the allowability of each of claims 26, 37 and 41 -- independently of any other claim(s) -- were detailed in Section VIII.B.3, above.

Independent claims 20 and 43 are nearly method/apparatus counterparts. However, claim 20 recites use of a natural language parser; claim 43 does not recite that the parser is a natural language parser.

Claim 36 is patentable independently of claim 30, from which it depends, by its recitation, *inter alia*, of using a computer to analyze a natural language corpus to identify a set of text segments therein, and to automatically discern a first collection of semantic relation structures therefrom.

X. CONCLUSION

The obviousness rejection is multiply-flawed, including:

- the rejection is premised on "admitted prior art" that isn't;
- Lenat repeatedly and adamantly castigated the approach followed by applicants as an unworkable "free lunch;" and
- the Office has not articulated any reason for the rejection of most of the claimed combinations.[39]

---

[39] As sole support for the final Section 103 rejection, the Office cited the Section 103 rejection expressed in Paper 4 (mailed 1/19/95). Paper 4 considered a different set of claims -- claims cancelled more than a year ago.

As such, reversal is required.

Respectfully submitted,

KLARQUIST SPARKMAN CAMPBELL
LEIGH & WHINSTON, LLP

By__________________
William Y. Conwell
Registration No. 31,943

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391

CLAIMS ON APPEAL

<Claims 16-19 allowed>

20. A method for generating a lexical knowledge base in a machine, said generating comprising the steps:

(a) using a natural language parser associated with said machine to parse a segment of text to obtain a logical form;

(b) using said machine to extract from the logical form a semantic relation structure, the semantic relation structure including at least a headword, a semantic relation, and a value;

(c) storing in a memory associated with said machine the semantic relation structure in association with the headword in the lexical knowledge base; and

(d) augmenting the lexical knowledge base by:

(1) inverting the semantic relation structure; and

(2) storing, in association with the value, the inverted semantic relation structure.

21. The method of claim 20 in which the semantic relation structure includes at least one element in addition to a headword, a semantic relation, and a value.

<Claims 22-23 objected to>

24. The method of claim 21 in which the parser is a broad coverage parser.

25. The method of claim 21 in which the parsing step includes:

(1) applying a first set of rules to the segment of text to yield a syntactic structure corresponding thereto; and

(2) applying a second set of rules to the syntactic structure to produce the logical form.

26. The method of claim 20 which includes:

providing a natural language corpus;

using the machine to analyze the corpus to identify a collection of text segments therein;

performing steps (a) - (d) a first time on the identified text segments to produce an augmented lexical knowledge base, the augmentation of the knowledge base serving to enhance subsequent text parsing; and

performing steps (a) - (d) a second time on the identified texts to further augment the knowledge base;

wherein the augmentation of the lexical knowledge base by the first performance of steps (c) and (d) enhances the parsing of the texts in the second performance of step (a).

27. The method of claim 26 which includes:

when performing step (b) the first time, extracting a first set of semantic relations; and

when performing step (b) the second time, identifying a second set of semantic relations, the second set of semantic relations being different from the first.

28. The method of claim 27 in which only the first set of semantic relations includes "hypernym."

29. The method of claim 27 in which only the first set of semantic relations includes "user," "domain" and "manner."

30. A machine-implemented method for generating a lexical knowledge base comprising the steps:

(A) using a computer to automatically process a text and identify at least a first semantic relation structure therein, said first semantic relation structure including at least four elements;

(B) using the computer to invert the first semantic relation structure to yield a second corresponding, inverted semantic relation structure; and

(C) storing data from the second semantic relation structure as part of a lexical knowledge base.

<Claims 31, 32 objected to>

33. The method of claim 30 in which the identifying step includes automatically parsing the text with a natural language parser, the parsing including:

applying a first set of rules to the text to yield a syntactic structure corresponding thereto;

applying a second set of rules to the syntactic structure to produce a corresponding logical form; and

extracting from the logical form said first semantic relation structure, said first semantic relation structure including at least a headword, a semantic relation, and a value.

34. The method of claim 33 in which the natural language parser analyzes for at least 8 of the following semantic relations: part, part_of, typical_subject, typical_subject_of, typical_object, typical_object_of, purpose, purpose_of, location_of, located_at, and synonym.

35. The method of claim 30 in which the identifying step includes string searching the text to discern semantic relation structures.

36. The method of claim 30 in which:

step (A) includes:

(a) providing a natural language corpus;

(b) analyzing the corpus by machine to identify a set of text segments therein;

(c) automatically discerning from the text segments a first collection of semantic relation structures; and

(d) storing data from the first collection of semantic relation structures as part of the lexical knowledge base;

step (B) includes:

(e) inverting the first collection of semantic relation structures to yield a second collection of corresponding, inverted semantic relation structures; and

step (C) includes:

(f) augmenting the lexical knowledge base by storing data from the second collection of corresponding, inverted semantic relation structures.

37. The method of claim 36 which includes performing steps (a) - (f) a first time, the augmentation of the knowledge base resulting from performance of step (f) serving to enhance subsequent discernment of semantic relation structures,

and then repeating steps (c) - (f) a second time; wherein augmentation of the knowledge base by the first performance of step (f) enhances the discerning of semantic relation structures by the second performance of step (c).

38. The method of claim 37 which includes discerning a first set of semantic relation structures in the first performance of step (c), and discerning a second set of semantic relation structures in the second performance of step (c), the second set of semantic relations being different from the first.

39. The method of claim 38 in which only the first set of semantic relations includes "hypernym."

40. The method of claim 38 in which only the first set of semantic relations includes "user," "domain" and "manner."

41. In a machine-implemented method of generating a lexical knowledge base comprised of using a computer to parse a collection of texts to identify semantic relation structures, and storing data from said semantic relation structures in a lexical knowledge base, an improvement comprising using the computer to iteratively parse the same collection of texts, successive parsings relying on successively enhanced versions of the lexical knowledge base, wherein semantic

relations identified in one parsing operation serve to enhance discernment of semantic relations in subsequent parsing operations.

42. The method of claim 41 in which the iterative parsing includes:

(1) applying a first set of rules to the collection of texts to yield a collection of syntactic structures corresponding thereto;

(2) applying a second set of rules to the syntactic structures to produce corresponding logical forms; and

(3) applying a third set of rules to the logical forms to obtain semantic relation structures, said structures including at least one instance of a triple comprising a headword, a semantic relation, and a value.

43. A system comprising:

(a) machine-implemented means for parsing a segment of text to obtain a logical form;

(b) machine-implemented means for extracting from the logical form a semantic relation structure, the semantic relation structure including at least a headword, a semantic relation, and a value;

(c) computer memory defining a lexical knowledge base in which the semantic relation structure associated with the headword is stored; and

(d) means for augmenting the lexical knowledge base including means for:

(1) inverting the semantic relation structure; and

(2) storing, in association with the value, the inverted semantic relation structure.

<Claim 43 objected to>

45. The apparatus of claim 43 in which the parsing means is a broad coverage parser.